Del panóptico al datapoint

Vigilancia y control por arquitectura

From the Panopticon to the Datapoint:
Surveillance and Control by Design

Mariano Enrique Torres Ponce

Abogado y Especialista en Derecho Informático

ORCID: 0009-0009-6284-3867

Marzo de 2022

Edición navegable y accesible. El contenido se mantiene sin modificaciones.

Resumen

El artículo propone una categoría para leer el presente tecnológico sin quedar atrapados en viejas metáforas ni en diagnósticos puramente declarativos. Llamamos poder operativo a la forma de autoridad que gobierna por diseño cuando el entorno técnico parametriza por adelantado el campo de lo posible, aprende de la conducta y legitima sus efectos por funcionamiento. La tesis sostiene que, frente a este desplazamiento desde la ley que habla hacia el sistema que actúa, un constitucionalismo eficaz debe traducir principios en propiedades verificables de diseño y operación.

A partir de esa premisa, el trabajo define garantías materiales que restituyen intervalo humano y control significativo mediante inteligibilidad práctica, trazabilidad suficiente, reversibilidad con costos acotados y proporcionalidad dinámica. Presenta además un método de verificación independiente integrado al ciclo de vida con auditoría contrafactual guiada por diseño, ensayos de fallo y estrés en escenarios reales, archivo peritable y métricas falsables como adherencia al default, gradiente de reversibilidad, opacidad efectiva y reducción de entropía comportamental.

La contribución se completa con un estándar institucional para compra pública y privada, separación de funciones, reconocimiento formal de la jurisdicción del default e interoperabilidad práctica con portabilidad que preserve contexto. El resultado ofrece una ruta para limitar el gobierno por optimización sin sofocar innovación y para reabrir el espacio del juicio en dominios donde la personalización y los valores por defecto gobiernan de hecho.

Palabras clave

poder operativo, constitucionalismo operacional, jurisdicción del default, inteligibilidad práctica, trazabilidad, reversibilidad, métricas falsables, auditoría contrafactual, arquitectura normativa.

Abstract

This article proposes a category to understand the technological present without remaining trapped in old metaphors or purely declarative diagnoses. We call operational power the form of authority that governs by design when the technical environment parameterizes the field of possibility in advance, learns from conduct, and derives legitimacy from its operation. The thesis argues that, facing this shift from law that speaks to systems that act, an effective constitutionalism must translate principles into verifiable properties of design and operation.

Based on this premise, the work defines material guarantees that restore space for human deliberation and meaningful control through practical intelligibility, sufficient traceability, reversibility with bounded costs, and dynamic proportionality. It also presents an independent verification method integrated into the life cycle with design-guided counterfactual auditing, stress and failure testing in real scenarios, auditable archives, and falsifiable metrics such as default adherence, reversibility gradient, effective opacity, and behavioral entropy reduction.

The contribution is completed with an institutional standard for public and private procurement, separation of functions, formal recognition of default jurisdiction, and practical interoperability with portability that preserves context. The result offers a path to limit governance by optimization without stifling innovation and to reopen the space for judgment in domains where personalization and default values govern in practice.

Keywords

operational power, operational constitutionalism, default jurisdiction, practical intelligibility, traceability, reversibility, falsifiable metrics, counterfactual auditing, normative architecture.

Resumen ejecutivo

Antecedentes: La vida cotidiana se organiza cada vez más en entornos digitales que deciden relevancias, rutas y ritmos sin reclamar autoridad visible. En ese escenario, la coacción explícita cede frente a la optimización y la legitimidad deja de apoyarse en el mandato para asentarse en la utilidad experimentada. Los sistemas técnicos configuran el espacio de acción con parámetros predeterminados y valores por defecto que gobiernan de facto, mientras el consentimiento se vuelve ritual y la personalización disuelve el conflicto en satisfacción de preferencias que el propio sistema ayudó a formar.

Vacío: La literatura crítica diagnostica esta transformación, pero carece de instrumentos operativos para limitarla. Las metáforas del panóptico y el extractivismo de datos describen síntomas sin fijar umbrales verificables ni remedios exigibles. El derecho permanece en registro declarativo cuando el poder contemporáneo actúa por realización técnica, lo que vuelve ineficaces las garantías que no se traducen en propiedades del diseño y la operación. Falta una categoría que distinga automatización asistente de gobierno por diseño y que permita someter la autoridad del entorno a contraste empírico con estándares falsables.

Propósito: El artículo propone una noción de poder operativo que aparece cuando confluyen parametrización previa del campo de acción legítimo mediante valores por defecto, aprendizaje por retroalimentación de la conducta que recalibra el entorno sin advertencia al usuario, y legitimación por funcionamiento que confunde fluidez con justificación. Solo en presencia de los tres rasgos hay gobierno por diseño con fuerza constituyente. Sobre esa base se articulan garantías materiales que restituyen intervalo humano y control significativo sin paralizar la innovación, traduciendo principios en condiciones verificables de inteligibilidad práctica, trazabilidad suficiente, reversibilidad con costos acotados y proporcionalidad dinámica calibrada por riesgo e impacto.

Metodología: Se presenta un método de verificación independiente integrado al ciclo de vida del sistema y no como control final. La auditoría contrafactual guiada modifica

de manera controlada valores por defecto, fricciones y ventanas de latencia para medir la fuerza de canalización de la arquitectura. Se ensayan escenarios de fallo, estrés y cambio de contexto para explorar puntos donde se anula la reversibilidad o el consentimiento queda vacío, con especial atención a poblaciones vulnerables y condiciones de borde que las validaciones promedio no detectan. El archivo peritable mínimo conserva versiones de modelos, rutas de datos, estados relevantes y cadena de custodia para que terceros puedan reconstruir decisiones controvertidas. Las métricas falsables convierten el diagnóstico filosófico en hipótesis de trabajo: adherencia al default para detectar jurisdicción ambiental del diseño, gradiente de reversibilidad para estimar el esfuerzo real de deshacer un resultado, opacidad efectiva como brecha entre lo declarado y lo reproducible, y reducción de entropía comportamental para advertir estrechamientos de trayectorias válidas sin ganancia proporcional en fines legítimos.

Resultados: La implementación institucional propuesta alinea incentivos y redistribuye capacidades de control. La compra pública y corporativa puede exigir desde el inicio archivo peritable, umbrales de activación para apagado seguro y degradación, así como cláusulas de verificación por terceros bajo confidencialidad. La arquitectura de gobernanza separa funciones entre quien diseña, quien opera y quien supervisa, mide intervención humana efectiva y activa un derecho a la latencia deliberativa que habilite pausas, escalados y desconexiones razonables. El reconocimiento formal de la jurisdicción del default asume que los valores por defecto gobiernan de hecho y por eso su activación demanda justificación reforzada, pruebas de no discriminación, opción alternativa simétrica y archivo de motivaciones. La interoperabilidad práctica y la portabilidad con equivalencia de desempeño evitan que la personalización se vuelva barrera de salida. El enfoque incluye una política de competencia en clave operacional con restricciones a la autopreferencia en sistemas de clasificación y relevancia, límites a integraciones que cierran el ecosistema y acceso razonable a interfaces y registros técnicos peritables para investigación y competencia.

Conclusión: La contribución principal es doble. Aporta un marco conceptual que diferencia automatización asistente de gobierno por optimización y un estándar operativo que permite pasar del principio a la práctica con pruebas y remedios. La novedad reside en fijar garantías materiales y métricas falsables que vuelven exigible el límite en sistemas que modulan conducta por arquitectura y no por mandatos explícitos. La legitimidad del entorno técnico no puede descansar solo en su buen funcionamiento, sino que debe sostenerse en límites verificables que preserven intervalo humano y control significativo. Cuando el sistema gobierna por optimización corresponde traducir garantías en requisitos de diseño, operación y prueba que cualquier tercero pueda auditar. Así el derecho vuelve a tener oficio en tiempo real, la innovación se ordena con responsabilidad y el poder operativo encuentra contención sin apagar la creatividad que

también necesitamos para diseñar libertades y no solo eficiencias.

Índice

Resumen **2**
Palabras clave 2

Abstract **3**
Keywords 3

Resumen ejecutivo **4**

1 Introducción. El poder que no se ve **8**

2 Filosofía del poder. De la ley al algoritmo **10**

3 Filosofía de la vigilancia. Del panóptico al datapoint **12**

4 Constitucionalismo operacional. Garantías materiales y umbrales de control **14**

5 Método de verificación independiente y ciclo de vida **17**

6 Economía política del poder operativo. Incentivos y captura del diseño **20**

7 Objeciones, límites y riesgos de segundo orden **23**

8 Implementación institucional. Estándar operativo para práctica pública y privada **26**

9 Conclusiones y agenda empírica **29**

Bibliografía **33**

Capítulo 1

Introducción. El poder que no se ve

El poder que organiza la vida contemporánea ya no necesita un rostro visible ni una sintaxis solemne. Opera como ambiente, como flujo que se confunde con lo cotidiano. Durante siglos la soberanía se exhibió en el cuerpo del gobernante y luego en la letra de la ley. Hoy la autoridad se redistribuye en arquitecturas técnicas que condicionan el espacio de acción con una eficacia tal que la utilidad percibida reemplaza a la obediencia como fuente de legitimidad (Foucault, 1975; Lessig, 1999; Bobbio, 1991).

Lo que antes se discutía en el lenguaje del deber ahora se resuelve en el diseño de los sistemas. Estos no argumentan ni persuaden, ejecutan sin pausa y producen efectos que ordenan silenciosamente la conducta del mismo modo que antaño lo hacía la gramática del castigo o la norma escrita, aunque con una potencia de modulación que desactiva el intervalo donde habitaban la interpretación, el juicio y el disenso, es decir el lugar de la libertad jurídica en sentido robusto.

Esta mudanza no equivale a una simple informatización del derecho. Expresa una transformación de la normatividad que desplaza el centro de gravedad desde el enunciado hacia la condición de posibilidad técnica. Donde el procedimiento habilitaba debate y demora, la interfaz promueve confirmación y continuidad con una suavidad que confunde experiencia de usuario con justicia y traduce preferencia en programación aprendida. La vigilancia ya no exige torre ni mirada frontal porque se integra en servicios personalizados, acumula rastros, predice trayectorias y recompone perfiles que devuelven al individuo una versión estadística de sí mismo, a tal punto que elegir se parece demasiado a confirmar la predicción optimizada para uno mismo, con la consiguiente erosión de la autonomía entendida como apertura de mundo y no solo como sensación de control subjetivo (Deleuze, 1990; Zuboff, 2019; Han, 2014).

Nombrar este fenómeno exige una categoría que no se agote en las metáforas del panóptico ni en el diagnóstico de la extracción opaca de datos, ya que el rasgo distintivo del presente no es únicamente la observación permanente sino la sustitución de la coacción explícita por la optimización como criterio de gobierno. Llamamos poder

operativo a la forma de autoridad que parametriza ex ante el campo de lo posible para que cumplir consista en seguir trayectorias ya dadas, aprende por retroalimentación de la conducta y apoya su legitimidad en la fluidez y la utilidad antes que en el mandato y la sanción. Esta diferencia clave permite trazar la frontera con la automatización instrumental y con la lógica de plataformas, que ya no actúa como mera infraestructura de servicios sino como principio operativo que define las condiciones mismas del orden (Rouvroy y Berns, 2013; Floridi, 2014).

Desde esta perspectiva se entiende por qué el derecho pierde primacía simbólica cuando permanece en registro declarativo, ya que el poder contemporáneo actúa por realización técnica y configura el entorno de acción con parámetros predeterminados y valores por defecto que gobiernan de facto. Cualquier propuesta garantista que pretenda ser efectiva debe traducir principios en condiciones operativas verificables. El primer movimiento del artículo fija esta hipótesis de desplazamiento uniendo genealogía y diagnóstico para mostrar que la ley que habla se volvió insuficiente frente al sistema que actúa, y que la tarea del pensamiento jurídico pasa por diseñar criterios exigibles de inteligibilidad, reversibilidad, trazabilidad y proporcionalidad capaces de reponer intervalo humano y control significativo allí donde la velocidad del cálculo tiende a clausurar toda demora, porque sin demora no hay deliberación y sin deliberación no hay libertad práctica (Lessig, 1999; Floridi, 2014; Foucault, 1975).

El hilo conductor que propongo sostiene que el constitucionalismo debe aprender a hablar el idioma de la ejecución sin renunciar a su vocación de límite. Esto exige distinguir con rigor cuándo un sistema gobierna por optimización y cuándo solo asiste decisiones humanas, además de establecer umbrales y remedios que vuelvan falsables tanto las descripciones críticas como las promesas técnicas. La introducción deja así planteado el programa conceptual que se desarrollará en las secciones siguientes, donde se precisará la demarcación del poder operativo, se articularán garantías materiales para un constitucionalismo operacional y se ofrecerán métodos de verificación y métricas que permitan someter a contraste empírico aquello que la filosofía del poder viene señalando desde hace décadas sobre la deriva desde la ley hacia el algoritmo, con el fin de reabrir el espacio del juicio dentro de un orden que tiende a confundirse con la inevitabilidad del funcionamiento (Deleuze, 1990; Winner, 1980).

Capítulo 2

Filosofía del poder. De la ley al algoritmo

La genealogía del poder permite leer un desplazamiento que pasa de la presencia soberana a la mediación textual y desde allí a la operación técnica. La ley organizó durante siglos el teatro del mandato en el registro del logos y sostuvo un régimen donde el tiempo de la deliberación hacía posible el juicio. El presente organiza la autoridad en arquitecturas que ejecutan y convierten la normatividad en condición de posibilidad, con lo cual el conflicto se traslada desde la discusión hacia el diseño y el cumplimiento se vuelve un efecto ambiental más que una decisión consciente del sujeto (Foucault, 1975; Deleuze, 1990).

Comprender este pasaje requiere asumir que la técnica no se reduce a un instrumento neutro y que los objetos técnicos tienen un modo de existencia propio que reconfigura memoria, tiempo y acción. La aceleración rompe el intervalo entre decisión y ejecución y borra el espacio donde antes habitaban la interpretación y la demora, de manera que la política se pliega a la logística y la razón práctica pierde espesor frente a la continuidad sin fisuras del flujo digital. Esto vuelve verosímil un poder que se legitima por funcionamiento y no por enunciación (Simondon, 1958; Stiegler, 2016; Virilio, 1995).

La diferencia específica del presente no radica solo en que el código regula. Importa que la regulación se desplaza desde la forma del deber hacia la forma del diseño, porque la interfaz propone trayectorias preferidas, el protocolo estandariza intercambios y los valores por defecto gobiernan sin presentarse como mandatos. Los artefactos condensan decisiones políticas y el derecho que permanece en el plano declarativo cede ante un orden que actúa por ejecución, confundiendo utilidad con legitimidad cuando lo que está en juego es la autoridad de hecho del entorno técnico sobre la conducta (Lessig, 1999; Winner, 1980).

Para evitar malentendidos conviene fijar una demarcación que distinga automatiza-

ción instrumental de poder constituyente. Llamamos poder operativo a la forma de gobierno que parametriza ex ante el campo de acción legítimo para que cumplir consista en seguir trayectorias ya dadas, reabsorbe la conducta como señal para recalibrar el entorno y sustituye la coacción por optimización y apoya su legitimidad en la fluidez y la utilidad. Allí donde no concurren estas tres condiciones solo hay poder asistido por técnica que comparte mecanismos, pero no alcanza fuerza normativa propia (Yeung, 2017; Pasquale, 2015; Zuboff, 2019).

La promesa de transparencia algorítmica exige pasar de declaraciones generales a reconstrucciones causales practicables por terceros, lo que obliga a enlazar filosofía del poder con teoría jurídica de la rendición de cuentas para que la autoridad del sistema no se escude en la caja negra ni en la retórica de la innovación inevitable (Nissenbaum, 2004; Hildebrandt, 2015).

El resultado pragmático de esta lectura propone un puente hacia la arquitectura institucional. Si la ley que habla se volvió insuficiente frente al sistema que actúa, el constitucionalismo debe aprender a expresarse en propiedades verificables del diseño y la operación, traduciendo límites en garantías materiales que restituyan intervalo humano, reintroduzcan comprensión efectiva y sometan a prueba la autoridad de los valores por defecto. El objetivo es reabrir el espacio del juicio sin sofocar la innovación y fijar un estándar que permita distinguir cuándo un sistema gobierna por optimización y cuándo solo asiste decisiones humanas con control significativo.

Capítulo 3

Filosofía de la vigilancia. Del panóptico al datapoint

La vigilancia que ofrecía una figura nítida en la torre central se volvió captura difusa de rastros y correlaciones. El punto de vista único cedió frente a la agregación ubicua de señales que reconstruyen perfiles dinámicos y gobiernan por probabilidad. El foco ya no está en ver a un individuo sino en anticipar trayectorias y modular contextos, de modo que la disciplina da paso a un control estadístico y este a una predicción que naturaliza su autoridad en la eficacia del acierto. El cálculo sustituye a la mirada y la ejecución técnica reemplaza a la orden que hablaba en nombre de la ley (Foucault, 1975; Deleuze, 1990; Zuboff, 2019).

La gobernanza algorítmica describe este pasaje con precisión porque actúa sin representar sujetos y opera sobre trazas que orientan la conducta por anticipación. El consentimiento se convierte en un ritual de aceptación que ya no garantiza comprensión efectiva, y la política del clic sustituye a la del pacto. La consecuencia más relevante para una teoría crítica del poder es que la coacción explícita se transforma en optimización persuasiva, y que la experiencia de personalización disuelve el conflicto en la satisfacción de preferencias que el propio sistema ayudó a formar. Se vuelve difícil distinguir deseo de diseño cuando la elección confirma una predicción preparada para nosotros (Rouvroy y Berns, 2013; Morozov, 2013).

Este régimen no existe sin la arquitectura de la red que concentra poder en plataformas capaces de fijar estándares de interacción, definir formas de visibilidad y establecer jerarquías de relevancia que deciden qué aparece y qué se oculta en la escena pública digital. La economía de nodos, enlaces y flujos tiene efectos normativos de hecho porque convierte el acceso en privilegio, el orden de aparición en criterio de verdad práctica y la opacidad de los sistemas de ranking en una fuente de autoridad que raramente se somete a deliberación democrática (Castells, 1996).

La privacidad deja de ser un recinto que se defiende por exclusión y se piensa como

integridad contextual que depende de expectativas situadas sobre circulación, finalidad y escala de los datos. En un entorno que recombina rastros para producir inferencias sobre identidades y riesgos, la vulneración no proviene solo de la revelación sino del cambio de contexto que altera significados y produce efectos sobre oportunidades, acceso y trato diferencial. Esto obliga a traducir garantías clásicas en condiciones materiales de uso, redistribución y auditoría de modelos y señales (Nissenbaum, 2004).

El vector que convierte la vigilancia en gobierno por diseño se advierte en la expansión de valores por defecto, empujes suaves y nudges algorítmicos que regulan por arquitectura. Cuando la plataforma ajusta la interfaz, modula fricciones y segmenta poblaciones con aprendizaje continuo, el empuje deja de ser general y adopta formas de personalización extrema. La autoridad se asienta en la utilidad experimentada y en la fluidez del recorrido, a tal punto que el valor por defecto funciona como una jurisdicción ambiental que orienta decisiones sin necesidad de mandato explícito y sin reconocer su fuerza normativa de hecho ante el derecho (Yeung, 2017; Thaler y Sunstein, 2008).

Desde aquí se entiende por qué la vigilancia contemporánea no es un apéndice de la seguridad sino parte del corazón operativo del orden social automatizado. El datapoint habilita una ingeniería fina de la conducta que exige deberes reforzados de inteligibilidad práctica y de trazabilidad causal, porque sin reconstrucción ex post de rutas de datos y pesos de señal la autoridad del entorno se vuelve irrecusable y la caja negra cancela la rendición de cuentas. La promesa de inteligibilidad técnica solo resulta relevante cuando se acopla a métodos y archivos peritables que permitan a terceros reproducir decisiones y verificar límites, ya que lo contrario reduce la transparencia a un gesto retórico (Kroll et al., 2017; Edwards y Veale, 2017; Ananny y Crawford, 2018).

La propuesta del constitucionalismo operacional conecta este diagnóstico con remedios verificables y retoma el hilo abierto en la sección anterior. Donde la vigilancia se confunde con personalización y la optimización sustituye a la coacción, se activan garantías materiales que restituyen intervalo humano y control significativo mediante latencia deliberativa, fricción justa y jurisdicción del default entendida como reconocimiento jurídico de la autoridad de los parámetros por defecto con deberes previos de explicación, opción alternativa simétrica y archivo para auditoría independiente. Solo así el derecho puede hablar el idioma de la ejecución sin abdicar de su vocación de límite y abrir un espacio donde el datapoint no clausure la deliberación pública, sino que quede sometido a umbrales y pruebas que vuelvan falsables tanto las descripciones críticas como las promesas de eficiencia (Lessig, 1999; Hildebrandt, 2015; Winner, 1980).

Capítulo 4

Constitucionalismo operacional. Garantías materiales y umbrales de control

Si el orden contemporáneo gobierna por diseño, un constitucionalismo a la altura del problema no se limita a proclamar principios, sino que los vuelve propiedades verificables de sistemas que actúan. La legitimidad ya no se juega en la forma del enunciado sino en la forma de la operación y por eso la promesa de derechos debe traducirse en condiciones técnicas que alguien pueda medir, auditar y refutar sin depender de la buena voluntad del operador ni de la opacidad del proveedor. El límite jurídico recupera eficacia allí donde la fluidez del entorno tiende a confundir utilidad con autoridad y funcionamiento con justificación normativa (Lessig, 1999; Winner, 1980; Hildebrandt, 2015).

Llamamos garantías materiales a esas condiciones mínimas de diseño y de operación que restituyen intervalo humano y control significativo sin paralizar la innovación. Se expresan en prácticas que pueden imponerse como obligaciones de resultado y de medio al mismo tiempo. La inteligibilidad práctica exige explicaciones que permitan reconstruir causas operativas y no solo narrativas tranquilizadoras. La trazabilidad requiere rutas de datos, versiones de modelos y decisiones registradas con cadena de custodia para peritaje independiente. La reversibilidad demanda opciones de corrección con costos acotados para el afectado y para el operador. La proporcionalidad dinámica impone que el grado de automatización se ajuste al riesgo y al impacto sobre derechos y que exista degradación segura cuando la incertidumbre crece. La opción alternativa simétrica garantiza que nadie quede atrapado en la solución única del proveedor cuando hay vías normativamente equivalentes que permiten cumplir la finalidad con menor intromisión o con mayor control social ex post (Edwards y Veale, 2017; Kroll et al., 2017; Nissenbaum, 2004).

Una noción decisiva en este marco es la jurisdicción del default. Los valores por defecto gobiernan de hecho y por eso deben someterse a reglas de justificación reforzada, con archivo de motivaciones, evidencia de pruebas de no discriminación en poblaciones relevantes y obligación de ofrecer caminos alternativos que no introduzcan fricciones punitivas. Allí donde el default concentra poder operativo, el estándar de control no puede reducirse a consentimientos rituales ni a políticas de privacidad ininteligibles. Necesita umbrales de activación, pruebas de estrés y mecanismos de auditoría que permitan a terceros replicar resultados y detectar sesgos o fallas de calibración que un operador complaciente podría pasar por alto, con especial atención a los efectos de acumulación que vuelven difícil atribuir responsabilidad cuando múltiples componentes interactúan en cadena bajo condiciones cambiantes (Yeung, 2017; Pasquale, 2015).

El control significativo se define entonces como la capacidad efectiva de intervenir antes, durante y después de la operación con medios proporcionados al riesgo. Esto supone tres planos coordinados. Un plano ex ante que fija requisitos de diseño y de pruebas con métricas de aceptabilidad que sean falsables y no meras declaraciones de intención. Un plano in itinere que habilita pausas, escalados y desconexiones razonables mediante latencia deliberativa y fricción justa. Un plano ex post que asegura reparación, revisión íntegra del expediente técnico y aprendizaje institucional que modifique parámetros y procedimientos cuando se comprueba daño o riesgo inaceptable. Sin esa tríada el gobierno por optimización tiende a consolidarse como destino y no como decisión sujeta a límites y razones públicas (Hildebrandt, 2015; Rouvroy y Berns, 2013).

Para volver exigible este programa se proponen umbrales operativos que articulan riesgo y automatización en tres zonas que orientan obligaciones y remedios. Una zona verde donde la automatización plena resulta admisible con auditorías periódicas y monitoreo continuo cuando el impacto es bajo y la reversibilidad es trivial. Una zona amarilla donde la automatización requiere control humano intensificado, opciones alternativas simétricas y pruebas de robustez con cobertura de subpoblaciones antes de cualquier despliegue a escala. Una zona roja donde la automatización debe ser limitada o prohibida hasta que existan garantías materiales capaces de absorber el riesgo residual con costos que no recaigan sobre el afectado ni sobre terceros vulnerables, y donde la degradación segura y el apagado por defecto sustituyen a la continuidad ciega del servicio cada vez que el sistema pierde calibración o deriva fuera de su dominio de validez (Edwards y Veale, 2017; Kroll et al., 2017).

La verificación deja de ser un apéndice burocrático cuando se expresa en métricas que cualquier autoridad independiente pueda calcular. Tasa de intervención humana efectiva que distinga supervisión nominal de intervención real. Tiempo a desacople

que mida cuánto demora un apagado seguro desde la detección de un desvío grave. Cobertura de trazas útiles que cuantifique cuántas decisiones pueden reconstruirse sin lagunas. Discrepancia por subpoblaciones que compare error y tratamiento entre grupos y exponga asimetrías. Índice de fricción justa que asegure que las vías alternativas no se vuelven caminos de castigo que desalientan legítimas elecciones. Registros de deriva que documenten cuándo y cómo cambia el comportamiento del sistema con datos nuevos o con cambios en el entorno. Todo ello en formatos estandarizados y con repositorios accesibles a peritos externos para que la transparencia no quede reducida a promesas comerciales o a portales decorativos sin datos peritables (Lessig, 1999; Winner, 1980).

El constitucionalismo operacional no se agota en técnicas de control, sino que propone una redistribución deliberada de capacidades. La arquitectura institucional debe separar funciones entre quien diseña, quien opera y quien supervisa, asegurar independencia material de los auditores, prever participación informada de colectivos afectados cuando el impacto sea estructural y crear incentivos públicos que premien diseños que incorporen límites como virtud y no como carga. Hay que evitar especialmente la captura regulatoria a través de estándares escritos por actores consolidados que afianzan posiciones dominantes bajo el ropaje de la seguridad o de la eficiencia inevitable, un riesgo bien documentado en otros ciclos tecnológicos y jurídicos que hoy reaparece con fuerza en el ecosistema algorítmico y de plataformas (Pasquale, 2015; Zuboff, 2019).

Este bloque configura el corazón propositivo del trabajo y abre paso a una metodología de implementación y prueba que permita pasar del principio a la práctica sin perder densidad jurídica ni realismo técnico. La sección siguiente desarrolla ese método en clave de verificación independiente y de ciclo de vida, con criterios de evidencia, procedimientos de auditoría y dispositivos de corrección que vinculan filosofía del poder, teoría del derecho y gobierno de sistemas en producción, para que el límite no sea un gesto simbólico sino una forma concreta de orientar la operación dentro de márgenes que mantengan abierto el espacio del juicio y de la deliberación pública.

Capítulo 5

Método de verificación independiente y ciclo de vida

El paso de los principios a las propiedades operativas exige un método que terceros puedan replicar sin depender de la confianza en el operador. La verificación se concibe como una práctica integrada al ciclo de vida del sistema y no como un trámite final, con instrumentos que vuelven falsables las promesas de inteligibilidad, trazabilidad, reversibilidad y proporcionalidad, y que permiten reconstruir causalmente lo ocurrido cuando un resultado es impugnado. La idea es que la autoridad técnica se someta a prueba del mismo modo en que el derecho somete a contraste la justificación de una decisión, con estándares que vinculan riesgo, automatización y remedios disponibles para la persona afectada (Kroll et al., 2017; Edwards y Veale, 2017).

La auditoría contrafactual guiada opera sobre parámetros de gobierno y no solo sobre datos. Se modifican de manera controlada valores por defecto, niveles de fricción y ventanas de latencia para estimar trayectorias alternativas y calcular su efecto normativo sobre decisiones y conductas. El foco está en la canalización por diseño y en su potencia para producir adhesión a rutas predeterminadas bajo apariencia de elección. El experimento no intenta reconstruir el modelo entero sino medir la fuerza de los empujes ambientales que orientan la acción y que muchas veces quedan fuera de las auditorías estadísticas tradicionales (Kroll et al., 2017; Ananny y Crawford, 2018).

Las pruebas adversariales sociotécnicas (red teaming) complementan esa mirada. El método consiste en someter intencionalmente el sistema a situaciones adversas para exponer dónde falla. Se prueban situaciones en las que las personas se equivocan, donde hay sobrecarga, donde cambian las condiciones o donde las señales no funcionan bien, prestando especial atención a grupos vulnerables y casos extremos que las pruebas normales no detectan. Lo que interesa no es solo encontrar sesgos, sino detectar dónde el sistema convierte la utilidad en autoridad y dónde la promesa de personalización erosiona la posibilidad de discrepancia con costos aceptables para el usuario y

para el operador (Edwards y Veale, 2017; Ananny y Crawford, 2018).

El archivo peritable mínimo asegura reconstrucción causal ex post. Se conservan versiones de modelos y políticas, estados de variables relevantes, semillas aleatorias, rutas de datos y cadena de custodia de registros técnicos en formatos aptos para auditoría externa, de modo que un tercero pueda reproducir el resultado controvertido y explorar alternativas cercanas sin depender de acceso privilegiado ni de explicaciones genéricas. Esta disciplina documental no es un lujo académico sino la condición para que la explicabilidad deje de ser retórica y se vuelva una práctica que habilite revisión, corrección y reparación en tiempo útil (Kroll et al., 2017; Edwards y Veale, 2017).

Las métricas falsables convierten el diagnóstico filosófico en hipótesis de trabajo. La adherencia al default mide qué proporción de decisiones sigue el valor por defecto cuando la salida impone costos no triviales y permite estimar la fuerza de canalización por diseño. El gradiente de reversibilidad estima el esfuerzo efectivo para deshacer un resultado en tiempo, recursos y pérdidas asociadas, y funciona como umbral para activar remedios robustos. La reducción de entropía comportamental captura el estrechamiento de trayectorias válidas a lo largo del tiempo sin ganancia proporcional en fines legítimos y alerta sobre sobre-gobierno operacional. La opacidad efectiva cuantifica la brecha entre lo que el sistema dice que hace y lo que puede reconstruirse causalmente después, y orienta deberes de transparencia pragmática más allá de promesas comerciales (Wachter et al., 2017; Ananny y Crawford, 2018; Winner, 1980).

El estándar de control se activa de manera escalonada. La falta de inteligibilidad mínima invalida cualquier presunción de proporcionalidad. La ausencia de reversibilidad sustantiva degrada la automatización a apoyo no vinculante y restituye la decisión al juicio humano con archivo reforzado. Una adherencia al default persistentemente alta por encima de umbrales pactados hace presumir canalización indebida y dispara deberes de explicación por adelantado, opción alternativa simétrica y trazabilidad suficiente para auditoría externa. La finalidad no es detener la innovación sino impedir que la optimización desplace derechos bajo apariencia de eficiencia (Kroll et al., 2017; Edwards y Veale, 2017).

El ciclo de vida integra estos elementos en tres momentos que dialogan entre sí. Ex ante se fijan requisitos de diseño y pruebas con métricas de aceptabilidad falsables, incluyendo cobertura por subpoblaciones y escenarios de degradación segura. In itinere se habilitan pausas razonables, escalados y desconexiones mediante latencia deliberativa y fricción justa que evitan la continuidad ciega. Ex post se garantiza reparación efectiva, revisión íntegra del expediente técnico y aprendizaje institucional que modifique parámetros y procedimientos cuando se constata daño o riesgo inaceptable. La graduación de obligaciones sigue la articulación entre riesgo e intensidad de automatización que se estableció como base del constitucionalismo operacional.

El método se ilustra con dominios donde el poder operativo muestra su fuerza constituyente. En prestaciones sociales sostenidas por sistemas de puntuación antifraude, pequeñas variaciones de umbral pueden suspender beneficios de forma automática y trasladar la carga probatoria al afectado, por lo que gradiente de reversibilidad y opacidad efectiva resultan críticos. En clasificación y visibilidad, las modulaciones de relevancia combinadas con señales de comportamiento funcionan como jurisdicción del default y justifican explicación por adelantado, opción alternativa y trazabilidad para auditoría. En triaje clínico asistido, el derecho a la latencia deliberativa preserva revisión humana en casos limítrofes y obliga a documentar intervenciones humanas correctivas con criterios clínicos verificables. En crédito al consumo con ofertas precargadas, una alta adherencia al default señala dominancia del diseño que amerita controles reforzados sobre proporcionalidad y entendimiento efectivo (Nissenbaum, 2004; Wachter et al., 2017; Pasquale, 2015).

El alcance y los límites se calibran con necesidad y proporcionalidad. El derecho a la latencia se modula en urgencias sin excluir revisión diferida, la fricción justa no habilita obstáculos artificiales ni burocratiza la experiencia, y la jurisdicción del default se reconoce solo frente a evidencia de canalización masiva, asimetrías de salida y ausencia de alternativas simétricas reales. La precisión sobre activación y límites blinda la propuesta frente a objeciones de ineficientismo y conserva margen para la experimentación responsable en diseño y operación, manteniendo abierto el espacio del juicio incluso en contextos de alta automatización (Thaler y Sunstein, 2008; Yeung, 2017; Mittelstadt et al., 2016).

Capítulo 6

Economía política del poder operativo. Incentivos y captura del diseño

El poder operativo no surge en el vacío. Crece donde la estructura de incentivos premia la fluidez aun cuando erosiona el intervalo del juicio y consolida posiciones dominantes mediante efectos de red, costos de cambio y aprendizaje acumulado que vuelven al proveedor casi inevitable. La utilidad experimentada se alinea con el interés del operador y la autoridad del entorno se legitima por funcionamiento más que por razón pública. Esto desplaza la competencia hacia la captura del diseño y convierte el default en un activo estratégico que gobierna de hecho sin reclamarlo en voz alta (Zuboff, 2019; Pasquale, 2015).

La acumulación de datos y de modelos produce rendimientos crecientes que favorecen el cierre progresivo del ecosistema. Cada interacción refuerza la asimetría informacional a favor del operador y transforma la personalización en barrera de salida, ya que cuanto más se adapta el sistema a un usuario y a un contexto, más costoso se vuelve abandonar la interfaz que organiza esa adaptación. Esto convierte a la portabilidad en un derecho formal con eficacia limitada si no se acompaña de condiciones técnicas que hagan trasladable la experiencia sin pérdida sustantiva de contexto y de desempeño real en el nuevo entorno (Stiegler, 2016; Nissenbaum, 2004).

Los mercados de atención y de relevancia exhiben una economía que premia la integración vertical del circuito de señales, clasificación y distribución. Esto genera estrategias de diseño que maximizan la adherencia al default y reducen la reversibilidad práctica, porque la continuidad del flujo incrementa el valor de la plataforma a ojos de terceros y refuerza la posición de intermediación como estándar de hecho. El operador puede imponer condiciones que trasladan costos de cumplimiento y de auditoría a los productores y a los usuarios sin que exista instancia eficaz de negociación simétrica, un fenómeno conocido en otros ciclos tecnológicos pero potenciado aquí por la opacidad técnica y por la velocidad de ajuste del sistema (Winner, 1980; Castells, 1996).

La retórica de la innovación inevitable funciona como cobertura cultural de esta economía política. Desplaza la discusión desde los límites a la promesa de eficiencia y pauta un horizonte donde toda restricción se presenta como freno anticompetitivo, aunque su función sea preservar el espacio del juicio. En ese marco los estándares de interoperabilidad suelen ser escritos por quienes tienen poder para capturar su letra y su cronograma de adopción, de modo que el resultado aparente de consenso consagra en los hechos las prácticas del operador dominante y vuelve costoso cualquier diseño alternativo que preserve derechos con fricción justa y latencia deliberativa (Pasquale, 2015; Zuboff, 2019).

Una política de competencia sensata en clave operacional no se agota en multas ni en remedios estructurales tardíos, sino que exige intervenir sobre los puntos donde el diseño cristaliza autoridad. La jurisdicción del default se vuelve una pieza regulatoria que reconoce el poder de los valores por defecto y somete su activación a justificación reforzada, pruebas de no discriminación en poblaciones relevantes, opción alternativa simétrica con costos razonables y archivo peritable que permita a terceros reconstruir cómo se eligió un parámetro y con qué evidencia. Sin este reconocimiento el control se disuelve en consentimientos rituales y portales de transparencia que dicen mucho, pero dejan intacta la operación (Yeung, 2017; Edwards y Veale, 2017).

La portabilidad con equivalencia práctica complementa esa línea. Trasladar datos sin trasladar contexto y comportamiento esperado equivale a mover cajas sin mudarse de casa. El derecho a salir con dignidad debe incorporar artefactos de migración que reconstituyan perfiles y preferencias con un margen razonable de desempeño en el nuevo entorno, junto con deberes de colaboración técnica mínimos entre operadores que no puedan ser bloqueados por cláusulas contractuales que invocan secreto comercial sin ofrecer alternativas verificables de auditoría independiente, un punto donde el equilibrio entre incentivos a la inversión y control social de la infraestructura común requiere precisión institucional y métricas falsables de éxito (Kroll et al., 2017; Wachter et al., 2017).

La adquisición de empresas emergentes y la integración de capas críticas de la arquitectura tecnológica operan como atajos para consolidar poder operativo porque acercan datos, modelos y canales de distribución bajo una misma gobernanza. Esto eleva el riesgo de que la experimentación responsable se convierta en entorno de pruebas permanente sin salida a reglas generales. Conviene activar umbrales que identifiquen concentraciones donde la línea entre automatización asistente y gobierno por optimización se cruza, con remedios que incluyan separaciones funcionales entre quien diseña, quien opera y quien decide estándares, obligaciones de acceso razonable a interfaces y registros técnicos peritables y límites a la autopreferencia en sistemas de clasificación y recomendación que deciden qué aparece y qué se oculta en la escena pública

digital (Pasquale, 2015; Castells, 1996).

El enfoque de constitucionalismo operacional ofrece criterios para orientar estas intervenciones sin sofocar emprendimiento ni investigación y desarrollo. No prohíbe automatizar en abstracto y no congela el diseño en una foto. Pide inteligibilidad práctica, trazabilidad suficiente, reversibilidad con costos acotados y proporcionalidad dinámica calibrada por riesgo e impacto, y somete a evidencia la eficacia de los remedios mediante métricas que cualquier tercero pueda calcular. Esto permite al regulador moverse con bisturí en lugar de martillo y distinguir dónde existe poder asistido por la técnica y dónde hay poder constituyente que gobierna por optimización y merece límites sustantivos en la operación cotidiana del sistema (Hildebrandt, 2015; Rouvroy y Berns, 2013).

El resultado es una economía política que vuelve visibles los lugares donde la utilidad encubre autoridad y donde la competencia aparente convive con una captura silenciosa del diseño. Habilita una agenda pragmática de control material del entorno sin negar la potencia creativa de la tecnología ni su aporte al bienestar. El objetivo no es ralentizar el progreso sino reabrir el espacio del juicio en un ecosistema que tiende a clausurarlo, combinando límites verificables con incentivos a soluciones que incorporen el límite como virtud y no como carga, de modo que la innovación ocurra con derechos dentro y no con derechos afuera (Winner, 1980; Lessig, 1999).

Capítulo 7

Objeciones, límites y riesgos de segundo orden

El constitucionalismo operacional enfrenta objeciones atendibles que merecen respuesta precisa, ya que una teoría que no soporta contraste riguroso se vuelve consigna. La primera objeción sostiene que convertir derechos en propiedades verificables encarece el diseño y puede enfriar la innovación. La réplica no confía en exhortaciones abstractas sino que propone proporcionalidad dinámica y métricas falsables para graduar obligaciones según riesgo e impacto, de modo que los costos regulatorios se concentren donde la automatización reemplaza juicio y no donde solo lo asiste. Este enfoque permite mover el límite con evidencia, someterlo a revisión y evitar que la retórica de la eficiencia desplace garantías que luego resultan mucho más costosas de reconstruir cuando el daño ya se produjo, especialmente en poblaciones expuestas o escenarios de alta asimetría informacional (Mittelstadt et al., 2016; Kroll et al., 2017; Edwards y Veale, 2017).

Una segunda objeción afirma que la explicabilidad es ilusoria en modelos complejos y que exigirla degrada el rendimiento. El estándar propuesto no confunde explicación con pedagogía universal, sino que pide inteligibilidad práctica para la decisión controvertida y archivo peritable que permita reconstrucción causal ex post por terceros, incluso si la arquitectura interna del modelo queda parcialmente oscurecida por razones legítimas de secreto o seguridad. La clave es pasar de promesas genéricas de transparencia a trazas, versiones y estados que habiliten verificación independiente, ya que sin esa disciplina documental la caja negra se convierte en escudo de impunidad y la apelación a la complejidad funciona como coartada cultural antes que como límite técnico insalvable (Wachter et al., 2017; Edwards y Veale, 2017; Ananny y Crawford, 2018).

La tercera objeción invoca secreto comercial y propiedad intelectual para bloquear auditorías. El marco operacional no desconoce ese interés, sino que lo reequilibra con

transparencia pragmática y verificación por terceros bajo deberes de confidencialidad, de modo que el titular proteja activos sin convertirlos en muro absoluto frente a la rendición de cuentas. Allí donde el diseño cristaliza autoridad y decide relevancias públicas, los remedios no pueden reducirse a la confianza en el operador, porque los artefactos condensan decisiones políticas y su neutralidad declarada rara vez resiste examen cuando se investigan dependencias, preferencias y efectos distributivos del entorno técnico (Pasquale, 2015; Winner, 1980).

Una cuarta objeción acusa paternalismo cuando se introduce fricción justa, latencia deliberativa o jurisdicción del default. El punto no es tutelar elecciones sino impedir que la reducción total de fricción anule comprensión efectiva y que valores por defecto gobiernen sin reconocer su fuerza normativa. La integridad contextual ayuda a precisar el daño cuando un mismo dato cambia de sentido según finalidad, escala o audiencia, y justifica intervenciones que devuelven control significativo sin burocratizar la experiencia ni imponer vueltas inútiles, ya que el objetivo es sostener libertad como práctica informada y no como ritual de clics que confirma predicciones ensambladas para cada quien (Thaler y Sunstein, 2008; Yeung, 2017; Nissenbaum, 2004).

Una quinta objeción advierte que el humano en el circuito se vuelve sello decorativo sin autoridad real. El estándar de control aborda esta deriva con umbrales de activación que degradan la automatización a apoyo no vinculante cuando faltan inteligibilidad mínima o reversibilidad sustantiva, y con un derecho a la latencia que no se limita a pausar sino que habilita escalado, desconexión razonable y archivo reforzado para responsabilizar a quien decide. La supervisión nominal cede ante la medición de intervención efectiva, que distingue presencia del humano de su capacidad real de torcer el curso cuando lo exige el caso, criterio que devuelve densidad al juicio práctico en contextos de alta automatización (Hildebrandt, 2015; Kroll et al., 2017).

Existen además riesgos de segundo orden que surgen de la propia adopción de métricas y estándares. La presión por mejorar indicadores puede inducir manipulación estratégica de indicadores y producir un efecto Goodhart por diseño, donde lo medido deja de representar bien lo que importa. La respuesta combina rotación de pruebas, auditorías aleatorizadas y pruebas adversariales sociotécnicas que explore rutas de evasión, junto con validación cruzada entre dominios y seguimiento de deriva para evitar que logros locales escondan degradaciones sistémicas. La literatura sobre equidad y aprendizaje automático muestra que los atajos aparecen en cuanto hay incentivos, por lo que la vigilancia de la medición es parte del método y no un epílogo burocrático que se delega a una lista de comprobación (Barocas et al., 2019; Ananny y Crawford, 2018).

Finalmente, puede objetarse que el marco es ambicioso para capacidades institucionales reales. Por eso la propuesta no pide perfección instantánea sino priorización por

riesgo, construcción incremental de pericia pública y estandarización mínima de repositorios, formatos y prácticas de archivo que bajen el costo de entrada de nuevas auditorías. El camino es modular y acumulativo, compatible con la experimentación responsable y con la incorporación de límites como virtud de diseño, a la vez que rechaza el destino de la innovación por inercia, porque no hay funcionamiento que justifique abdicar del juicio cuando lo que está en juego es la autoridad de hecho del entorno técnico sobre la conducta colectiva (Winner, 1980; Lessig, 1999).

Capítulo 8

Implementación institucional. Estándar operativo para práctica pública y privada

Traducir el marco en práctica exige mover el centro de gravedad de la política tecnológica hacia condiciones verificables de diseño y operación, con reglas que alineen incentivos y redistribuyan capacidades de control. El núcleo es simple y exigente a la vez. Pide que la autoridad del entorno técnico se someta a prueba del mismo modo en que el derecho contrasta razones, lo que implica colocar en el contrato, en la compra pública y en la gobernanza interna requisitos de inteligibilidad práctica, trazabilidad suficiente, reversibilidad con costos acotados y proporcionalidad dinámica calibrada por riesgo e impacto, de manera que la eficacia no oculte la justificación y que la utilidad no sustituya al límite. El resultado no es un manual de buenas intenciones sino un estándar operativo capaz de ser auditado por terceros que no dependen de la confianza en el operador ni del acceso discrecional que concede el proveedor.

La compra tecnológica pública y corporativa es el primer tornillo que se puede girar con efectos sistémicos, ya que la demanda puede exigir archivo peritable mínimo, versiones de modelos, rutas de datos, semillas y estados relevantes, junto con umbrales de activación para el apagado seguro y la degradación cuando crece la incertidumbre. No se trata de abrir la caja negra sin límites, sino de asegurar transparencia pragmática que habilite reconstrucción causal y verificación independiente bajo deberes de confidencialidad, porque el secreto comercial protege inversiones sin borrar la rendición de cuentas cuando el artefacto decide relevancias públicas. Incluir estas cláusulas en contratos y pliegos convierte al constitucionalismo operacional en criterio de elegibilidad y no en adorno retórico.

El diseño institucional debe separar funciones entre quien construye, quien opera y quien supervisa, con independencia material y técnica de los auditores y con parti-

cipación informada de colectivos afectados cuando el impacto sea estructural. Esta arquitectura distribuye el poder de interpretación y reduce el riesgo de sello decorativo del humano en el circuito. Fija umbrales de activación que degradan la automatización a apoyo no vinculante cuando faltan inteligibilidad mínima o reversibilidad sustantiva, y mide intervención efectiva en lugar de conformarse con presencia simbólica. La supervisión deja de vivir en el papel cuando existen pausas configurables, escalados, derecho a la latencia deliberativa y posibilidad real de torcer el curso operativo sin penalidades ocultas.

La gobernanza de plataformas y servicios intensivos en datos requiere además reconocer jurisdicción del default como categoría regulatoria, porque los valores por defecto gobiernan de hecho. Su activación demanda justificación reforzada, pruebas de no discriminación en poblaciones relevantes, opción alternativa simétrica con costos razonables y archivo que permita reproducir cómo se eligieron parámetros y con qué evidencia, con especial atención a efectos de acumulación que complejizan la atribución de responsabilidades cuando interactúan múltiples componentes. Sin este reconocimiento formal, el control se disuelve en consentimientos rituales que simulan libertad sin garantizar comprensión efectiva ni vías reales de salida.

Para evitar la captura del diseño por actores consolidados que redactan estándares a su medida, conviene combinar tres dispositivos de defensa. Primero, métricas falsables y públicas que cualquier tercero pueda calcular, como adherencia al default, gradiente de reversibilidad, reducción de entropía comportamental y opacidad efectiva, con rotación de pruebas y auditorías aleatorizadas a fin de desactivar el efecto Goodhart. Segundo, entornos de prueba con salida obligatoria a reglas generales y con criterios de éxito previamente definidos, que impidan convertir la excepción en régimen permanente. Tercero, interoperabilidad práctica que no se agote en el traslado de datos sino que preserve contexto y desempeño razonable al migrar, para que la personalización no se vuelva barrera de salida ni moneda de captura.

El derecho procesal puede acompañar esta mudanza con reglas probatorias que distribuyan cargas según asimetrías reales. Una adherencia al default persistentemente alta por encima de umbrales pactados puede generar presunción de canalización indebida y activar deberes de explicación por adelantado y de oferta de alternativas, mientras la ausencia de archivo peritable mínimo deshace la presunción de proporcionalidad y traslada al operador la carga de mostrar que su solución no anuló la reversibilidad sustantiva. Esta gramática incentiva diseños que incorporan límites como virtud y desalienta la opacidad eficiente que externaliza costos sobre usuarios y terceros vulnerables.

La política de competencia en clave operacional completa el cuadro con remedios sobre los puntos donde cristaliza autoridad. Separaciones funcionales entre diseño,

operación y definición de estándares, restricciones a la autopreferencia en sistemas de clasificación y relevancia, acceso razonable a interfaces y registros técnicos peritables para quienes compiten o investigan, y control de adquisiciones que integran capas críticas de la arquitectura tecnológica cuando elevan el riesgo de cierre del ecosistema. El objetivo no es castigar tamaño sino prevenir que la utilidad experimentada se alinee inercialmente con el interés del operador hasta convertir el default en jurisdicción privada sobre la escena pública digital.

Este estándar operativo no congela la innovación, sino que le da forma y responsabilidad. Al amarrar funcionamiento con justificación y utilidad con límites, reabre el espacio del juicio dentro de procesos de alta automatización y ofrece una ruta incremental para reguladores, operadores y tribunales. La agenda empírica inmediata consiste en validar umbrales y remedios en dominios como prestaciones sociales, clasificación y moderación, triaje clínico y crédito al consumo, con comparación cruzada de costos de reversión, cobertura de trazas útiles y tiempos de desacople, de modo que la teoría del poder operativo se someta a evidencia y pueda traducirse en práctica estable sin sofocar la creatividad técnica que, bien orientada, también puede diseñar libertades.

Capítulo 9

Conclusiones y agenda empírica

El recorrido mostró que el poder operativo no es un simple efecto de informatización sino una forma de autoridad que gobierna por diseño y que traslada el conflicto desde el lenguaje del deber hacia la arquitectura de la acción. Su eficacia se legitima en la utilidad experimentada y en la fluidez del funcionamiento, de modo que la ley que habla pierde primacía cuando no logra traducirse en propiedades verificables del sistema que actúa y cuando el intervalo del juicio queda absorbido por la continuidad del flujo técnico que organiza opciones, ritmos y salidas posibles para cada quien en función de perfiles y trayectorias aprendidas en tiempo real. El propio entorno captura y recombina señales de comportamiento sin fricción relevante para el usuario común que confirma predicciones con su propia elección mediada por la interfaz que lo hospeda y lo conduce sin violencia visible, aunque con autoridad de hecho sobre su recorrido cotidiano (Foucault, 1975; Deleuze, 1990; Lessig, 1999; Zuboff, 2019).

El nombre que aquí se fijó para describir este fenómeno permite distinguir entre automatización asistente y gobierno por optimización. Solo en el segundo caso aparecen los tres rasgos que confieren fuerza constituyente al entorno: parametrización previa del campo de lo posible con valores por defecto que canalizan, aprendizaje por retroalimentación de la conducta que recalibra el contexto y legitimación por funcionamiento que confunde eficiencia con justificación. Esta diferencia justifica límites materiales capaces de reabrir el espacio del juicio mediante inteligibilidad práctica, trazabilidad suficiente, reversibilidad con costos acotados y proporcionalidad dinámica que degrade la automatización cuando el riesgo crece o cuando la incertidumbre rompe el dominio de validez del sistema. Sin esa degradación la promesa de control humano se vuelve sello decorativo y la caja negra opera como escudo de impunidad técnica frente a la rendición de cuentas que el derecho espera sostener con razones públicas y con pruebas peritables por terceros independientes en contextos de conflicto real y no en demostraciones de laboratorio con condiciones ideales de iluminación y de carga (Hildebrandt, 2015; Kroll et al., 2017; Edwards y Veale, 2017).

El constitucionalismo operacional propuesto ofrece una gramática que une filosofía del poder y teoría del derecho con método de verificación que no se agota en buenas prácticas, sino que se ancla en mediciones falsables que cualquier autoridad ajena al operador pueda calcular sin favores de acceso ni permisos discrecionales. La jurisdicción del default aparece como categoría decisiva junto con el derecho a la latencia deliberativa y la fricción justa, ya que ahí se juega la posibilidad de intervenir antes, durante y después de la operación, de activar apagados seguros, de abrir rutas alternativas que no castiguen con demoras punitivas y de reconstruir causalmente lo ocurrido cuando una persona impugna un resultado que la afecta en su acceso a derechos o en su visibilidad pública. Esto ocurre dentro de ecosistemas donde la clasificación decide oportunidades y el orden de aparición funciona como verdad práctica para millones de casos por día sin deliberación expresa ni control significativo de las asimetrías que el sistema refuerza con su propio éxito operativo en mercados de atención y de relevancia que premian la continuidad y la adherencia a la ruta preferida por el diseño de la plataforma dominante del sector (Winner, 1980; Yeung, 2017; Castells, 1996).

La agenda empírica inmediata exige pasar de la retórica a los números con estudios de campo que midan adherencia al default en dominios sensibles, gradiente de reversibilidad efectiva con tiempos, costos y pérdidas asociadas para el afectado y para el operador, opacidad efectiva como brecha entre lo que se declara y lo que puede reconstruirse después con archivo peritable mínimo, y reducción de entropía comportamental como indicador de estrechamiento de trayectorias válidas sin ganancia proporcional en fines legítimos. Todo ello con cobertura por subpoblaciones que permita detectar asimetrías de error y de trato que se diluyen en promedios y que vuelven inaceptable cualquier presunción de proporcionalidad cuando la divergencia supera umbrales pactados con publicidad suficiente. La rotación de pruebas mitiga el efecto de manipulación estratégica de indicadores y de Goodhart que inevitablemente aparece cuando la métrica empieza a importar para reputación y para elegibilidad en compras públicas o corporativas de gran escala que fijan estándares de hecho más rápido que cualquier ley escrita en un parlamento que todavía piensa en plazos anuales para cambios que la operación introduce a diario en ciclos incrementales que jamás se anuncian como cambios regulatorios aunque lo son en su eficacia real sobre la conducta (Ananny y Crawford, 2018; Wachter et al., 2017; Barocas et al., 2019).

La agenda institucional acompaña esa línea con requisitos contractuales y procesales que bajen a tierra la verificación independiente. La compra pública y la compra corporativa pueden exigir versiones de modelos, rutas de datos, estados de variables relevantes, semillas y registros técnicos con cadena de custodia para auditoría externa bajo deberes de confidencialidad que protejan secreto comercial sin bloquear la reconstrucción causal cuando hay conflicto. El proceso judicial o administrativo debe recono-

cer presunciones razonables que distribuyan cargas de prueba según asimetrías reales, por ejemplo, activar una presunción de canalización indebida cuando la adherencia al default supere rangos esperables para poblaciones comparables con igual finalidad y con alternativas simétricas disponibles en teoría, pero no en la práctica de la interfaz que impone fricciones que desalientan la salida. Estas fricciones merecen el nombre de jurisdicción ambiental de la plataforma sobre el usuario que no tiene poder de negociación salvo el derecho a decir no en abstracto mientras en los hechos dice sí porque la vida le urge y el sistema le ofrece la ruta más corta al resultado que necesita en ese momento, aunque el precio sea quedar atrapado en un ecosistema que cierra puertas de salida con la misma sutileza con que abre ventanas de entrada (Lessig, 1999; Pasquale, 2015).

El programa de investigación comparada suma un plano transnacional donde interoperabilidad práctica y portabilidad con equivalencia de desempeño funcionen como antídotos frente a la captura del diseño por actores consolidados que redactan estándares a su medida. Mover datos sin mover contexto no garantiza libertad de salida y solo traslada el problema de una caja a otra. Conviene diseñar artefactos de migración que reconstituyan perfiles con un margen razonable de degradación controlada y con métricas de éxito previamente publicadas, junto con umbrales de concentración que disparen remedios sobre capas críticas de la arquitectura tecnológica cuando la integración vertical de señales, clasificación y distribución eleve el riesgo de cierre y vuelva inverificables las rutas por donde circulan las decisiones que afectan derechos y oportunidades a escala poblacional. Los impactos desiguales que un promedio niega quedan revelados cuando la estadística desagregada muestra subpoblaciones y condiciones de borde y no solo el centro del campo donde el sistema luce bien porque fue calibrado con datos que lo favorecen por construcción histórica de su propio dominio (Castells, 1996; Zuboff, 2019).

El itinerario propuesto mantiene un principio de prudencia que evita el maximalismo normativo y el laissez-faire tecnoutópico. Combina zonas de automatización admisible con auditorías periódicas cuando el impacto es bajo, zonas donde la intervención humana reforzada y la opción alternativa simétrica se vuelven obligatorias, y zonas donde la automatización debe degradarse o prohibirse hasta que existan garantías materiales capaces de absorber el riesgo residual sin trasladarlo a quien menos puede soportarlo. Este enfoque no pretende congelar la innovación sino atarla a límites que ensanchan libertad práctica y que devuelven densidad al juicio en un mundo donde la continuidad sin pausa tiende a confundir rapidez con razón y funcionamiento con justicia.

La teoría aquí elaborada intenta despejar esa confusión con herramientas que admiten prueba y que no dependen de la buena voluntad del operador ni del encanto

cultural de la promesa de eficiencia inevitable que domina el discurso público sobre tecnología desde hace demasiado tiempo sin someterse a contrastes que la igualen ante el derecho y ante la evidencia de su propia operación. Los sistemas complejos merecen límites que sepan hablar su idioma y a la vez recordarle que no hay cálculo que agote la vida ni optimización que valga como coartada cuando lo que se pierde es el intervalo mínimo donde la deliberación respira y el desacuerdo encuentra un lugar para existir. Una plataforma que todo lo mide puede dejar fuera aquello que más importa para sostener una vida libre y justa en común con otros bajo leyes que limitan y habilitan a la vez, porque ese es su oficio cuando no renuncian a la tarea de gobernar la operación con reglas que se pueden ver y que se pueden discutir con pruebas en la mano y no solo con promesas comerciales o con fe en la ingeniería de turno.

El desafío pasa por construir instituciones que sobrevivan a la moda y al ciclo de versión de una herramienta cualquiera que promete mucho y rinde bastante, aunque nunca tanto como sus promotores quieren hacernos creer si no hay límite que obligue a mostrar números y a aceptar correcciones cuando la realidad las exige. Gobernar la técnica pide tiempo, método y coraje cívico para decir hasta acá y también para decir sigamos, pero con condiciones, con archivo y con control significativo que vuelva a poner al juicio en su sitio sin expulsar a la creatividad de la sala donde también se la necesita para diseñar libertades y no solo eficiencias. El operador que entiende el poder de los valores por defecto y lo usa hasta el final necesita reglas que se puedan medir y remedios que se puedan hacer valer en la práctica y no solo en los papeles de un mundo ideal que no existe fuera de los folletos.

Bibliografía

Ananny, M., y Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society, 20*(3), 973–989. https://doi.org/10.1177/1461444816676645

Barocas, S., Hardt, M., y Narayanan, A. (2019). *Fairness and machine learning*. https://fairmlbook.org/

Bobbio, N. (1991). *El futuro de la democracia*. Fondo de Cultura Económica. https://fce.com.ar/tienda/politica/el-futuro-de-la-democracia/

Castells, M. (1996). *The rise of the network society*. Blackwell.

Deleuze, G. (1990). Post-scriptum sur les sociétés de contrôle. *L'Autre Journal, 1*, 3–7.

Edwards, L., y Veale, M. (2017). Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. *Duke Law & Technology Review, 16*(1), 18–84. https://scholarship.law.duke.edu/dltr/vol16/iss1/2/

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press. https://global.oup.com/academic/product/the-fourth-revolution-9780199606726

Foucault, M. (1975). *Surveiller et punir: Naissance de la prison*. Gallimard.

Han, B.-C. (2014). *Psicopolítica: Neoliberalismo y nuevas técnicas de poder*. Herder.

Hildebrandt, M. (2015). *Smart technologies and the end(s) of law: Novel entanglements of law and technology*. Edward Elgar Publishing. https://doi.org/10.4337/9781849808774

Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., y Yu, H. (2017). Accountable algorithms. *University of Pennsylvania Law Review, 165*(3), 633–705. https://pennlawreview.com/2017/02/23/

accountable-algorithms/

Lessig, L. (1999). *Code and other laws of cyberspace*. Basic Books.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., y Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2), 1–21. https://doi.org/10.1177/2053951716679679

Morozov, E. (2013). *To save everything, click here*. PublicAffairs.

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review, 79*(1), 119–158. https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10/

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press. https://www.hup.harvard.edu/books/9780674970847

Rouvroy, A., y Berns, T. (2013). Gouvernementalité algorithmique et perspectives d'émancipation: Le disparate comme condition d'individuation par la relation? *Réseaux, 177*(1), 163–196. https://doi.org/10.3917/res.177.0163

Simondon, G. (1958). *Du mode d'existence des objets techniques*. Aubier.

Stiegler, B. (2016). *Dans la disruption: Comment ne pas devenir fou?* Les Liens qui Libèrent.

Thaler, R. H., y Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press. https://yalebooks.yale.edu/book/9780300262285/nudge/

Virilio, P. (1995). *La vitesse de libération*. Galilée.

Wachter, S., Mittelstadt, B., y Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law, 7*(2), 76–99. https://doi.org/10.1093/idpl/ipx005

Winner, L. (1980). Do artifacts have politics? *Daedalus, 109*(1), 121–136. https://www.jstor.org/stable/20024652

Yeung, K. (2017). 'Hypernudge': Big data as a mode of regulation by design. *Information, Communication & Society, 20*(1), 118–136. https://doi.org/10.1080/1369118X.2016.1186713

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.